NAME: Jinyi Xia
STUDENT ID: 2021212057
CLASS NUMBER: 2021211802

ASSIGNMENT 1

1 Exercise 1

When compiling the given statement, 7 tokens will be generated. The 1st token is the keyword `int`, the 2nd token is the identifier `a3`, the 3rd token is the operator `=`, the 4th token is the identifier `a`, the 5th token is the operator `*`, the 6th token is the integer literal 3, and the 7th token is the delimiter `;`.

2 Exercise 2

In a string of length n, there are

1. $n+1$ prefixes;

2. $n-1$ proper prefixes;

3. 1 prefix of length m;

4. 1 suffix of length m;

5. 1 (if $m \neq n$) or 0 (if $m = n$) proper prefix of length m;

6. $\dfrac{n(n+1)}{2}+1$ substrings;

7. 2^n subsequences.

3 Exercise 3

1. $\left((\varepsilon|a)^* b^*\right)^*$ denotes the language of all strings over the alphabet $\{a,b\}$.

2. $(a|b)^* a\,(a|b)\,(a|b)$ denotes the language of all strings over the alphabet $\{a,b\}$ that end with *aaa*, *aab*, *aba* or *abb*.

3. $a^* b a^* b a^* b a^*$ denotes the language of all strings over the alphabet $\{a,b\}$ where character b appears exactly 3 times.

4 Exercise 4

1. 86-0755-$[1-9][0-9]^7$;

2. $a\,(a|b)^* b$;

3. Let *consonant* $\rightarrow (b|c|d|f|g|h|j|k|l|m|n|p|q|r|s|t|v|w|x|y|z)^*$, the regular expression is
consonant **a** *consonant* **e** *consonant* **i** *consonant* **o** *consonant* **u** *consonant*.

5 Optional Exercise 1

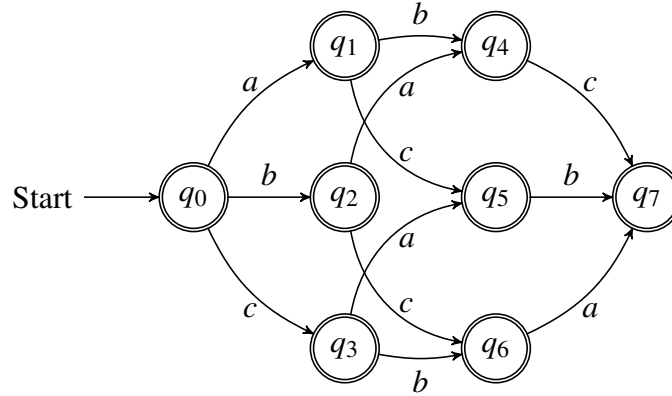The DFA for the language of all strings over Σ without repeated letters is shown in fig. 1.

Figure 1: The DFA's transition diagram

According to the transition diagram, the regular expression for the language of all strings over Σ without repeated letters is

$$(\varepsilon|a)\,(\varepsilon|b)\,(\varepsilon|c)\,|$$
$$(\varepsilon|a)\,(\varepsilon|c)\,(\varepsilon|b)\,|$$
$$(\varepsilon|b)\,(\varepsilon|a)\,(\varepsilon|c)\,|$$
$$(\varepsilon|b)\,(\varepsilon|c)\,(\varepsilon|a)\,|$$
$$(\varepsilon|c)\,(\varepsilon|a)\,(\varepsilon|b)\,|$$
$$(\varepsilon|c)\,(\varepsilon|b)\,(\varepsilon|a)\,.$$